2011-06

Examining Work Performance in Immersive Virtual Environments versus Face-to-Face Physical Environments through Laboratory Experimentation

Examining Work Performance in Immersive Virtual Environments versus Face-to-Face Physical Environments through Laboratory Experimentation

Richard Bergin
Naval Post Graduate School
Monterey, CA
rdbergin@nps.edu

Bryan Hudgens
Naval Post Graduate School
Monterey, CA
bjhudgen@nps.edu

Mark Nissen PhD
Naval Post Graduate School
Monterey, CA
mnissen@nps.edu

Abstract

With increasing proliferation of virtual environments for serious work as well as play, we are confronted by new challenges pertaining to how such environments can be leveraged to promote and induce effective work performance. In terms of organization and management, we are challenged further by decisions regarding which work activities to perform in virtual environments versus their physical counterparts; our level of understanding remains relatively primitive, and the literature remains divided. The research described in this article examines work performance in virtual versus physical environments through laboratory experimentation. We construct an immersive virtual environment, in which people interact via avatars to perform collaborative work that matches a physical environment where the same collaborative work is performed. Exercising great care to match experiment conditions and control for factors other than work environment, results elucidate important comparative performance effects and suggest compelling follow-on experiments as well as practical implications.

1. Introduction

With the increasing proliferation of virtual environments for serious work as well as play, we are confronted by new challenges pertaining to how such environments can be leveraged to promote and induce effective work performance [1]. Following considerable maturation of the Web and the advent of Web 2.0, businesses, government agencies, military units, non-profit groups and other organizations are working increasingly in virtual environments as well as—and in many cases instead of—their physical counterparts [2].

Additionally, as technology continues to advance, virtual environments are becoming increasingly immersive in terms of user experiences [2]. In most cases, immersiveness in virtual environments is designed to replicate the experiences of their physical counterparts (e.g., office buildings, group meetings, familiar tools and other work artifacts) closely, but in many other cases, designers are exploiting experiences (e.g., co-presence, teleportation, morphing) that are possible only through virtual environments [3].

In terms of organization and management, we are challenged further by decisions regarding which work activities to perform in virtual environments versus their physical counterparts. Unfortunately for the organizational leader and manager, our level of understanding in this area remains relatively primitive, and unfortunately for the researcher, the literature remains divided in terms of guidance [4].

This article describes a pilot study to examine work performance in virtual versus physical environments through laboratory experimentation. We construct an immersive virtual environment, in which people interact via personalized avatars to perform collaborative work that matches a physical environment where the same collaborative work is performed. Exercising great care to match experiment conditions and control for factors other than work environment, results elucidate important comparative performance effects and suggest compelling follow-on experiments as well as practical implications.

In the balance of the paper, we draw from the educational psychology, immersiveness, and media richness literatures to motivate a set of research hypotheses to address the virtual-physical question well. We then detail our research design and report in turn the key findings and results. The paper closes with a set of conclusions, recommendations for practice, and topics for future research along the lines of this study.

2. Background

In this section we summarize succinctly a core set of literature from educational psychology, immersiveness, and media richness literatures.

1530-1605 2011

2.1. Educational psychology

The question of what can be accomplished via virtual versus physical environments has been addressed for many years by researchers in education, particularly where one or more information technologies (e.g., videotaping, TV broadcasting with telephone, video teleconferencing, web-based instruction) are used to enable distance education beyond the classroom. Historically there has been a strong bias toward classroom teaching (i.e., physical environment) and against distance education (i.e., virtual environment). Indeed, several governing bodies of higher education (e.g., accreditation boards) have issued blanket bans of distance education [5], and where standards and policy organizations (e.g., Western Cooperative for Educational Telecommunication, Higher Education and Policy Council of the American Teachers Federation, Institute for Higher Education Policy) have sought to bolster distance education through quality standards, they have tended to focus on minimum standards to equate distance with classroom education [6].

However, numerous studies [7] have compared the efficacy of classroom versus distance education, and most cases show no significant differences [8]. Indeed, Bates and Poole [5] indicate that "… the research evidence indicates clearly that technology-based teaching can be just as effective as face-to-face teaching" (p. 19). They go further by noting how technology enables some teaching techniques that are infeasible in the classroom, and they suggest that in some respects distance education can be better than its classroom counterpart (p. 23).

This complements the group decision support systems literature [9], which has shown for two decades that some aspects of group performance (e.g., mitigating rank and status differences, overcoming shyness and language difficulties, developing higher quality work products, examining a more complete range of alternatives and perspectives) are indeed better when people's interactions pertaining to information work are mediated technologically than in face-to-face interactions. This leads us to our first research hypothesis.

Hypothesis 1a. Performance of information work activities through highly immersive virtual environments will exceed that of the same activities performed through physical environments.

2.2. Immersiveness

Slater and Wilbur [10] suggest that "immersion and the resulting level of presence experienced by an individual can be predicted though quantifying various display characteristic such as the degree to which stimuli from the real world are excluded from the user, the number of sensory modalities accommodated by the system, how panoramic the displays are, and the resolution of the displays" and by matching task characteristics and individual preferences to sensory information types such as auditory or visual. Slater, Linakis, Usoh, and Kooper [11] later suggest that higher levels of presence better facilitate real-world behaviors in a virtual environment, and when those behaviors are beneficial, task performance increases. Building on Slater and Wilber [10] Bystrom, Barfield, and Hendrix [12] develop the Immersion, Presence, Performance (IPP) model that seeks to describe "the effects of display technology, task demands, and attentional resource allocation on immersion, presence, and performance in virtual environments." The IPP model provides a link between the level of immersion, as defined by various display characteristics and spatial fidelity, presence, and task performance.

Research on the influence of presence on task performance in virtual environments has generated mixed findings. Stanney, Mourant, and Kennedy [13] suggest that human performance in virtual environments is positively influenced by the level of presence provided by a virtual world. Slater, Sadagic, Usoh, and Schroeder [14] when studying puzzle solving skills in a virtual environment discovered a positive relationship between presence and group performance. McGill and Anand [15] explored the influence of various levels of media vividness on apartment selection performance and discovered that vivid stimuli unrelated to the primary task may not increase involvement and distract the decision maker from their task. Witner [16] discovered that "there is a weak but consistent positive relationship between presence and task performance", but that individual differences influence the level of presence experienced, and these differences may not be influenced by the level of immersiveness provided by a particular display.

These mixed findings on the influence of presence in a virtual environment on task performance appear to result from the matching of task characteristics to information sensory requirements. Bystrom, et al. [12] suggest that "whether the display technology used to produce virtual environments will assist task performance will be dependent on the type of task to be performed and the level of immersion.

Current research suggests that particular types of tasks might benefit from higher levels of presence in highly immersive virtual environment. This leads us to posit that depending on task characteristics; highly immersive virtual environments will increase performance.

Moreover, as technology advances and virtual environments become increasingly immersive (e.g., online video games and social networking sites, virtual and augmented reality, telepresence; see [17-20]), the range of feasible learning and experiential effects continues to expand. This leads us to our next research hypothesis.

H1b. Performance of information work activities through virtual environments will increase in proportion with the degree of immersion.

2.3. Media richness

Daft & Lengel [21] define media richness "as the ability of information to change understanding within a time interval. Communication transactions that can overcome different frames of reference or clarify ambiguous issues to change understanding in a timely manner are considered rich. Communications that require a long time to enable understanding or that cannot overcome different perspectives are lower in richness. In a sense, richness pertains to the learning capacity of a communication." Oral media such as face-to-face and the telephone are considered to have higher levels of media richness than written media, such as interoffice mail. Additionally, synchronous media such as telephone and live chat sessions are considered to have higher levels of media richness than asynchronous media such as interoffice email or electronic mail [22].

Media richness theory rests on the assumption that organizations process information to reduce uncertainty and equivocality [21]. Uncertainty is defined by Galbraith [23] as "the difference between the amount of information required to perform the task and the amount of information already possessed by the organization." Equivocality is the existence of multiple and conflicting interpretations about an organizational situation, or confusion, disagreement and lack of understanding about a particular problem-solving event [24-26]. Considering media richness theory as a prescriptive model [21, 27], high and low levels of media richness provide distinct advantages in terms of reducing either equivocality or uncertainty.

Daft & Lengel [27] conclude that written or text based media that are low in media richness are preferred for unequivocal messages, while a face-to-face environment that is high in media richness is preferred for messages containing equivocality. Rice and Shook [28] suggest that media low in richness, such as business letters, convey less of a social presence and are less effective in terms of reducing equivocality through bargaining, negotiation, and conflict resolution. They also suggest that media low in richness are able to reduce uncertainty through the exchange of facts and information.

Within the context of business, government, and military operations, the level of uncertainty, equivocality and the resultant information processing needs vary depending on the operational environment. The comparative efficacy of using communication media within a highly immersive virtual environment that is low in media richness versus face-to-face interaction that is high in media richness remains somewhat unclear and controversial, however.

Indeed, the prescriptions from Educational Psychology above (especially that performance in virtual environments will exceed that in physical counterparts) conflict in part with those from Media Richness Theory (especially that performance in media-rich environments will exceed that in media-poor counterparts). This is the case in particular where virtual environments (e.g., with non-immersive, textual interfaces) are expected to reflect low media richness. Moreover, this question and theoretical conflict take on increasing importance now, as many businesses, government agencies and military organizations are turning increasingly toward network-centric operations, through which most interactions between people are mediated by relatively non-immersive (esp. textual), media-poor environments. This leads to a conflicting hypothesis that can be tested empirically.

Hypothesis 2. *Performance of information work activities under conditions of high media richness will exceed that of the same activities performed under conditions of low media richness.*

3. Research design

In this section, we draw heavily from [29, 30, 30] to summarize the research design used in this pilot study. Building upon prior experimentation, we employ the ELICIT multiplayer intelligence game in a laboratory setting to conduct a series of experiments and examine how a highly immersive virtual environment versus a face-to-face physical environment affects performance in the context of an information work task environment. We begin by describing this task environment and then outline the participants, groups, protocols, controls, manipulations and measurements used for experimentation in the virtual environment case. We close this section with a summary of the physical environment case.

3.1. Task environment

The task requires a team of participants performing information work through the roles of intelligence analysts to collaborate and identify a fictitious and stylized terrorist plot. The fictitious terrorist plot is described through a set of informational clues called “factoids” that have been developed systematically. The game’s design is similar to the Parker Brothers’ board game *Clue* in that it requires each participant to analyze clues and combine information from other participants to identify key aspects of the fictitious plot. Each factoid describes some aspect of the plot, but none is sufficient to answer all of the pertinent questions (i.e., Who will execute the attack? What is the target to be attacked? Where will the attack take place? When will the attack take place?).

The factoid distribution is designed so that no single participant can solve the problem individually and that the team of participants cannot solve the problem until after the final distribution. In other words, the participants must collaborate to solve the problem, and they are required to do so for a minimum of ten minutes. Evidence from previous experiments [31] suggests that play requires substantially more time (e.g., an hour or more).

Participants play the game in one of two modes. 1) For the highly-immersive virtual environment, each participant manipulates his or her avatar within Second Life (SL) . Each participant shares and posts factoids with other participants via avatars seated around a virtual table within the SL virtual room representation.

Each participant has access to a set of five functions inside the SL room (i.e., List, Post, Pull and Share factoids; Identify plot details). The virtual environment captures time-stamped interactions (e.g., Post, Pull, Identify, List functions) including, for instance, when and which factoids are distributed to each participant, when and which common area are viewed by each participant, when and which factoids are shared between each participants, and the time stamped results of each participant’s Identify attempt (i.e., to identify the who, what, where and when).

2) For the physical environment, they play through face-to-face interaction in real-world rooms equipped with physical tables and white boards. Factoids are time stamped and distributed to the participants on pieces of paper, and paper “postcards” are time stamped and used to collect participants’ Identify attempts.

The game requires considerable cognitive and collaborative effort to play well (i.e., identify the pertinent details of a terrorist plot), but experience indicates that such effort is within the capabilities of many people and groups.

3.2. Participants

Participants for this study are comprised of PhD and Master’s students, staff, and faculty volunteers in the information sciences and other departments at a major US university. All participants have undergraduate college degrees, and most have completed graduate work at the masters and/or PhD levels.

In this experiment, participants are organized hierarchically. Such organization stratifies them into three functional levels. The ELICT game is played with 14 participants who assume the following pre-assigned roles: A Senior Leader (executive) is responsible for the organization as a whole and has four Team Leaders (middle managers) reporting directly. Each team leader in turn has two or three Team Members (Operators) reporting directly and is responsible for one set of details associated with the terrorist plot. For instance, Team Leader (Who) and his or her team are responsible for the “who” details (e.g., which terrorist organization is involved) of the plot, Team Leader (What) and his or her team are responsible for the “what” details (e.g., what the likely target is), and so forth for “where” and “when.”

3.3. Treatment groups

To address concerns with learning and bias, participants are pre-assigned randomly to the different roles in each game. Additionally, different yet structurally equivalent versions of the game are played each time, and the two experiment sessions are conducted four months apart; this reduces learning opportunities yet ensures that both problem-solving and information-sharing tasks are comparable. We summarize protocols and manipulations for each group below.

3.3.1. Highly immersive virtual environment.

Protocols and manipulations for the SL environment are designed to be consistent with most prior experiments using ELICIT. One week prior to the experiment, participants are instructed on how to set up an SL account and how to create an individual avatar. Participants report to a networked classroom on their assigned day for the experiment. Once seated, participants are allotted 40 minutes to spend time learning how to maneuver their avatars in the SL environment, to interface with the five ELICIT functions, and to read a set of instructions pertaining to both the experiment and the ELICIT environment; they are encouraged to ask questions about the experimental settings and environment. Once participants read the instructions they have ten minutes to discuss their approach to the problem-solving scenario with others on their team and take a short break before beginning the game. The first five minutes are allotted to discussion between the Senior Leader and four Team Leaders. The next five minutes are allotted to separate discussions between each Team Leader and his or her three team members.

Once the game begins, each participant receives unique factoids in three phases: 1) two factoids initially when the game begins, 2) one after five minutes, and 3) one at the ten-minute mark. Role-specific factoids are distributed automatically by ELICIT and in a manner ensuring: a) that no participant can solve the plot alone, and b) that the plot cannot be solved until all factoids have been distributed. Factoids are time stamped and appear automatically on participants' ELICIT screens, after which they can be posted, pulled and shared. As noted above, all Post, Pull, Share and Identify actions are time stamped and logged by the game server.

Participants do not communicate with one another during game play, only the ELICIT functions (esp. Post, Pull and Share) are readily available—no face-to-face verbal communication is allowed.

Participants are informed explicitly of one another's assigned pseudonyms and their self-selected avatar names. This helps to match conditions in the physical environment where participants' identities and pseudonyms are known.

Additionally, the SL virtual rooms limit each participant's Post (i.e., sharing factoids with others) and Pull (accessing factoids posted by others) access to his or her specifically assigned virtual room. Those participants in the "who" team, for instance, are allowed to Post to and Pull from only their virtual room (i.e., the "who" room) noted above. The only exception applies to the Senior Leader, who has post-pull access to all four virtual rooms. This is comparable to prior experiments and matches the physical counterpart to this experiment condition.

Further, in this experiment team members are permitted to share factoids only with members of their own teams, and Team Leaders are permitted to share only with members of their respective teams in addition to the Senior Leader. Moreover, the Senior Leader communicates only with Team Leaders. This reinforces hierarchical communication and chain of command, and it and matches the physical counterpart to this experiment condition.

The simulation ends after approximately 50 minutes. All participants are given the option to identify the plot details if they have not done so already. Participants are instructed to Identify only once during game play, and they are incentivized both to solve the plot individually as well as to collaborate so that others on the team (esp. the Senior Leader) solve the plot quickly and accurately.

3.3.2. Physical environment.

Protocols and manipulations for the physical environment are designed to be as consistent as possible with those summarized above for the highly immersive virtual environment. Participants report to a real-world classroom with physical desks, tables and whiteboards on their assigned day for the experiment. As above, once seated, participants are allotted ten minutes to read a set of instructions pertaining to both the experiment and the physical environment; they are encouraged to ask questions about the experimental settings and environment. Once participants read the instructions they have ten minutes (as above) to discuss their approach to the problem-solving scenario with others in their group and take a short break before beginning the game.

Once the game begins, each participant receives the same unique factoids in the same three phases summarized above. In this case, however, role-specific factoids are distributed manually on pieces of paper. The time of distribution is noted on each factoid. Unlike in the virtual environment above, we make no attempt in this physical environment to record or note the time when factoids are posted (e.g., on the white board) pulled (e.g., by a participant taking notes from the white board) or shared (e.g., verbally). However, consistent with the virtual environment above, the time of each Identify action is recorded.

In great contrast with the virtual environment, participants communicate with one another during game play using only face-to-face conversation and the white board; no computer-mediated communication is allowed.

As above, each participant is pre-assigned randomly a pseudonym associated with his or her role in the game, and as noted above, participants are informed explicitly of one another's pseudonyms. Likewise, participants' physical separation in different rooms limits their Post (i.e., sharing factoids with others) and Pull (accessing factoids posted by others) access to specific white boards in each of the rooms within this manipulation. Those participants in the "who" group, for instance, are allowed to Post to and Pull from only of the physical white board in the "who" room. The only exception applies to the Senior Leader, who has access to information and white boards in all four rooms. This preserves the limitations imposed in the virtual environment above.

Also as above, in this experiment team members are permitted to share only with members of their own teams, and Team Leaders are permitted to share only with members of their respective teams in addition to the Senior Leader. Further, the Senior Leader communicates only with Team Leaders. This reinforces hierarchical communication and chain of command.

When the simulation ends, participants are given the option to identify the plot details if they have not done so already. Participants are instructed to Identify only once during game play, and they are incentivized both to solve the plot individually as well as to collaborate so that others on the team (esp. the Senior Leader) solve the plot quickly and accurately.

3.4. Measurements

Following Leweling & Nissen [29] we operationalize performance as a two-dimensional dependent variable comprised of: 1) speed (i.e., time to identify plot details correctly) and 2) accuracy (i.e., correct identification of plot details). These dependent measures are informed by literature in the psychological and organizational domains that suggest a trade-off exists between time and accuracy in tasks requiring high cognition and/or advanced motor skills [32-37] at both the individual and team/group levels of analysis.

In the first component, speed pertains to how long it takes a participant to submit his or her identification of the terrorist plot details. For ease of comparison, the scale for this speed measurement is normalized to a 0-1 scale, with 1 being more desirable (i.e., faster). Measuring and normalizing time is straightforward, as the time for each participant's identification is logged to the nearest second by the software. Specifically, each participant's elapsed time is recorded when he or she uses ELICIT to Identify the plot. To construct a scale in which faster speeds (i.e., shorter times to Identify) result in larger values, a baseline time is established as the maximum time required for the slowest of all participants (i.e., 3365 seconds in this experiment). Each participant's time to identify is related to this baseline and normalized to produce a scaled score according to the formula: speed = (3365 – time) /3365; that is, an individual participant's time (say, for example, 2300 seconds) would be converted to a speed score as: speed = (3365 – 2300) / 3365 = 0.3165. All participants' times are converted to speed scores in this same manner and using this same baseline.

The second component of performance, accuracy, refers to the quality of the identification of the impending terrorist attack (i.e., Who, What, Where, and When). Each participant's Identify action is scored with a value of 1 for a correct answer to the Who, What and Where aspect of the solution. Because the When aspect of the solution includes three components (i.e., Month, Day, and Time), each participant's Identify action is scored with a value of 1/3 for a correct answer in this category (i.e., getting all three When elements right earns 1 point). The resulting sum is divided by four to construct a [0-1] scale; that is, an individual participant's Identify (say, for example, identifies the Who, What and Where aspects correctly but is correct only on the day and not the month or time components of the When aspect) would be converted to an accuracy score as: accuracy = (1 + 1 + 1 + 1/3) / 4 = 0.83. We note here that this research design does not enable us to test Hypothesis 1a in our pilot study. Results from testing this hypothesis are reserved for the follow-on study.

4. Results

In this section, we summarize results from the experiment, beginning with an overview of key statistics and followed by discussion of their implications.

Table 1. Descriptive statistics

Variable	Mean	N	Standard Deviation	Standard Error
ID Time HIVE (in Seconds)	2972	13	269.1	74.6
ID Time FTF (in Seconds)	**2554**	14	279.2	74.6
Who Score HIVE	.62	13	.506	.140
Who Score FTF	**.79**	14	.426	.114
What Score HIVE	.38	13	.506	.140
What Score FTF	**.79**	14	.426	.114
Where Score HIVE	.08	13	.277	.077
Where Score FTF	**.79**	14	.426	.114
When Score HIVE	.23	13	.160	.044
When Score FTF	**.33**	14	.320	.086
Identify Composite HIVE	.33	13	.251	.070
Identify Composite FTF	**.67**	14	.344	.092

4.1. Statistical results

Table 1 presents the descriptive statistics. In the table, "HIVE" refers to highly immersive virtual environment, and "FTF" refers to face-to-face physical environment. To enhance interpretation, we use bold font to highlight the higher performing score in each HIVE-FTF contrast. Speed is faster (i.e., less time to Identify) in the FTF condition.

Table 2 presents results of the hypotheses tests. We employ the basic t-test of independent samples to examine differences in means between HIVE and FTF results.

Each row of the table summarizes a contrast between HIVE and FTF. For instance, the first row summarizes the difference between identification time (in seconds) for the HIVE and FTF sessions (i.e., HIVE time minus FTF time). We report this difference to remind the reader which score is "better"; in all cases except for the identification time (longer identification time signifies slower speed), a positive difference indicates the HIVE version is "better", whereas a negative difference indicates the FTF version is "better". In the case of identification time, the opposite is true since longer time indicates "worse" performance.

For instance, in the first row we report the mean difference (418 seconds), t value (3.953), degrees of freedom (25) and significance (.001, independent sample t-test).

Table 2. Results of hypotheses tests

Variable	Statistical Support Mean Value	t	df	Significance (2-tailed)
ID Time HIVE - ID Time FTF in Seconds	418	3.953	25	.001
Who Score HIVE - Who Score FTF	-.170	-.949	25	.352
What Score HIVE - What Score FTF	-.401	-2.234	25	.035
Where Score HIVE - Where Score FTF	-.709	-5.160	22.5	.000
When Score HIVE - When Score FTF	-.103	-1.040	25	.308
HIVE Composite – FTF Composite	-.346	-2.965	25	.007

The difference in identification time is significant (p = .001). The accuracy contrasts for correctly identifying "who" and "when" are not significant (p = .352 and p = .308, respectively); however, the accuracy contrasts are significant for identifying "what" (p=.035) and "where" (p =.000). Additionally, the accuracy contrast for the composite identification score is significant (p=.007). In summary, our results suggest that the FTF approach leads to significantly faster and more accurate performance overall.

4.2. Key findings

The results summarized above contain two important findings for organizational leaders seeking to better understand the comparative performance of people working in media-rich face-to-face environments versus highly immersive virtual counterparts. The results of this study support media richness theory, while at the same time they provide a plausible explanation for the theoretical conflict between the Educational Psychology literature and media richness theory.

As summarized above, performance in the highly immersive virtual environment is worse than in the physical environment. This provides support *against* hypotheses H1a. Highly immersive virtual environments that are low in media richness, such as the one used in this experiment and those commonly found in businesses, government agencies, military units, non-profit groups, and other organizations, provide little opportunity for professionals to reduce task equivocality, often resulting in reduced group accuracy. Although highly immersive virtual environments can provide a sense of place and copresence, provide access to experts, and accelerate information sharing among distributed decision makers, thus reducing uncertainty, such access does not appear to be compelling in this experiment.

Alternatively, in terms of time and accuracy, performance in the media rich face-to-face environment is better than in a highly immersive virtual environment. This provides support *for* hypothesis H2.

In this experiment, both speed and accuracy are sensitive to physical or virtual environments. This calls into question what appears to be a strong assumption reflected in organizations: that physically distributed, virtual environments should be employed. Such environments do not appear to gain either a speed or accuracy advantage. This suggests that leaders and policy makers may benefit by rethinking their organizational assumptions, particularly where speed and accuracy are important.

Further, considering this experiment, some of the theoretical conflicts between the educational psychology and media richness literatures may be resolved by exploring the variation in task uncertainty and task equivocality among tasks. Educational tasks focus on particular sets of learning objectives and tend to be static, well defined, and unequivocal. Business, government, and military tasks related to sense and response activities within an emergent environment are more dynamic, however, often requiring higher levels of interpretation and negotiation to resolve equivocal goals and objectives.

Within an educational setting where task equivocality is low, it therefore makes sense how virtual environments that are low in media richness can outperform face-to-face environments by providing efficient real-time access to expert knowledge and information that reduces uncertainty in the task environment. We gain further insight into how business, government, and military organizations, operating within emergent, sense-and-respond ecospaces where task equivocality is high, can benefit more from face-to-face environments than virtual environments through greater shared and more accurate interpretation of the task environment. Hence a plausible resolution for the theoretical conflict between the Educational Psychology literature and media richness theory resides in defining the task in terms of the level of task uncertainty and task equivocality.

5. Conclusion

With increasing proliferation of virtual environments for serious work as well as play, we are confronted by new challenges pertaining to how such environments can be leveraged to promote and induce effective work performance. In terms of organization and management, we are challenged further by decisions regarding which work activities to perform in virtual environments versus their physical counterparts; our level of understanding remains relatively primitive, and the literature remains divided.

The research described in this article examines work performance in virtual versus physical environments through laboratory experimentation. We construct an immersive virtual environment, in which people interact via avatars to perform collaborative work that matches a physical environment where the same collaborative work is performed. Exercising great care to match experiment conditions and control for factors other than work environment, results elucidate important comparative performance effects and suggest compelling follow-on experiments as well as practical implications.

Specifically, participants in the physical environment outperform counterparts in the virtual environment. Hence—where accuracy and timing are concerned—business, government, and military leaders and policy makers may find cause to rethink the widespread organizational assumption that favors physically distributed, virtual work.

Further, through consideration of how uncertainty and equivocality affect the kinds of environments associated generally with educational contexts versus those encountered by business, government, and military in a dynamic sense and respond environment, we offer some resolution of the theoretical conflict between Educational Psychology and Media Richness Theory: where uncertainty prevails (e.g., in the educational context), virtual environments may be adequate and even outperform their physical counterparts, but where equivocality predominates (e.g., business, government, and military processes), richer media offered through physical environments appear to be important and to support superior performance.

These results suggest compelling follow-on experiments as well as practical implications. For instance, one could study other organizational forms as alternatives to the Hierarchy. As another instance, one could expand the range of capabilities available to participants (esp. via avatars) within the virtual environment and examine how different degrees of immersiveness, for example, affect performance. As a third instance, software agents can be developed to play the ELICIT game—either in conjunction with or in lieu of human participants—across a range of virtual environments and organizational forms.

In terms of practical implications, these results call into question the predominate, physically distributed, virtual environment employed for work. Leaders and policy makers may have cause to rethink their organizing assumptions, particularly where decision making accuracy and timing is critical for responding to emergent business, government, and military events. Although much information is collected from distributed sources—and hence must be physically distributed by necessity—the *analysis* of such information does not have the same necessary cause for physically distributed work, and leaders and policy makers may find it useful to collocate analysts in physical environments that afford media-rich, face-to-face interactions.

As with any study, the series of experiments reported in this article have limitations. For instance, we are unable to examine the influence of copresence in a highly immersive virtual environment or contrast alternate organizational forms. The literature suggests that such variables are important; hence our results should be interpreted as partial until the complementary experiments noted above have been completed. Also, as noted above, we place this paper in the context of a pilot study and plan to examine H1b along with a full set of hypotheses through follow-on work.

Although the research described in this article provides only one metaphorical step in this direction, it is an important step, one that challenges prevailing organizing wisdom in terms of distributed collaborative information work and decision making, and one that offers to reconcile theoretical conflict pertaining to performance in physical versus virtual environments. We look forward to continuing along the lines of this investigation, and we welcome others to join and complement our work.

6. References

[1] K. J. McLennan, *The Virtual World of Work : How to Gain Competitive Advantage through the Virtual Workplace.* Charlotte, NC: Information Age Pub., 2008.

[2] N. Yee, "The Labor of Fun: How Video Games Blur the Boundaries of Work and Play," vol. 1, pp. 68, 2006.

[3] E. Castronova, *Synthetic Worlds: The Business and Culture of Online Games.* Chicago, IL: University of Chicago Press, 2005.

[4] A. W. Bates and G. Poole, *Effective Teaching with Technology in Higher Education: Foundations for Success.* San Francisco, CA: Jossey-Bass, 2003.

[5] D. S. Alberts, *the Unintended Consequences of Information Age Technologies.* Washington, DC: CCRP, 1996.

[6] D. S. Alberts, Garstka, Hayes and Signori, *Understanding Information Age Warfare.* Washington, DC: CCRP, 2001.

[7] D. S. Alberts, Garstka and Stein, *Network Centric Warfare.* Washington, DC: CCRP, 1999.

[8] D. S. Alberts and R. E. Hayes, *Command Arrangements for Peace Operations.* Washington DC: CCRP, 1995.

[9] D. S. Alberts and J. Moffat, "Maturity levels for NATO NEC command," UK Defence Science & Technology Laboratory, UK, 2006.

[10] M. Slater and S. Wilbur, "Through the looking glass world of presence: A framework for immersive virtual environments," in *FIVE '95 Framework for Immersive Virtual Environments,* 1995, .

[11] M. Slater, V. Linakis, M. Usoh and R. Kooper. Immersion, presence, and performance in virtual environments: An experiment using tri-dimensional chess. Presented at Virtual Reality and Software and Technology Conference.

[12] K. E. Bystrom, "A conceptual model of the sense of presence in virtual environments," *Presence,* vol. 8, pp. 241, 1999.

[13] K. M. Stanney, "Human factors issues in virtual environments: A review of the literature," *Presence,* vol. 7, pp. 327, 1998.

[14] M. Slater, "Small-group behavior in a virtual and real environment: A comparative study," *Presence,* vol. 9, pp. 37, 2000.

[15] A. L. McGill, "The effect of vivid attributes on the evaluation of alternatives: The role of differential attention and cognitive elaboration," *The Journal of Consumer Research,* vol. 16, pp. 188, 1989.

[16] B. G. Witmer, "Measuring presence in virtual environments: A presence questionnaire," *Presence,* vol. 7, pp. 225, 1998.

[17] K. C. Allard, "Somalia Operations: Lessons Learned," 1995.

[18] A. W. Bates and G. Poole, *Effective Teaching with Technology in Higher Education: Foundations for Success.* San Francisco, CA: Jossey-Bass, 2003.

[19] M. D. P. Bolger, "Command or Control," *Mil. Rev.,* pp. 69-79, 1990.

[20] W. Broll, J. Ohlenburg, I. Lindt, I. Herbst and A. Braun, "Meeting technology challenges of pervasive augmented reality games," in 2006, .

[21] R. L. Daft and R. H. Lengel, "Organizational Information Requirements, Media Richness and Structural Design," *Management Science,* vol. 32, pp. 554-571, May, 1986.

[22] M. L. Markus, "Electronic Mail As the Medium of Managerial Choice," *Organization Science,* vol. 5, pp. 502-527, Nov., 1994.

[23] J. R. Galbraith. (1977, *Organization Design* .

[24] R. L. Daft and N. B. Macintosh, "A Tentative Exploration into the Amount and Equivocality of Information Processing in Organizational Work Units," *Adm. Sci. Q.,* vol. 26, pp. 207-224, Jun., 1981.

[25] L. K. Trevino, "Message equivocality, media selection, and manager performance: Implications for information systems," *MIS Quarterly,* vol. 17, pp. 355, 1987.

[26] K. E. Weick, *The Social Psychology of Organizing.* Addison-Wesley, 1979.

[27] R. L. Daft and R. H. Lengel. (1984, Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior 6*pp. 191-233.

[28] R. E. Rice and D. E. Shook, "Relationships of Jop Categories and Organizational Levels to Use of Communication Channels, Including Electronic Mail: A Meta-Analysis and Extension," *The Journal of Management Studies,* vol. 27, pp. 195, Mar, 1990.

[29] T. A. Leweling and M. E. Nissen, "Hypothesis testing of edge organizations: Laboratory experimentation using the ELICIT multiplayer intelligence game," in *Proceedings International*

Command & Control Research & Technology Symposium, Newport, RI, 2007, .

[30] E. H. Powley and M. E. Nissen, "Trust-mistrust as a design contingency: Laboratory experimentation in a counterterrorism context," in Washington, DC, 2009, .

[31] Parity Communications Inc., "Experiments in command and control within edge organizations: Final report," Boston, MA, Tech. Rep. EBR06-0002, 2006.

[32] B. Beersma, J. R. Hollenbeck, S. E. Humphrey, H. Moon, D. E. Conlon and D. R. Ilgen, "Cooperation, Competition, and Team Performance: Toward a Contingency Approach," *Academy of Management Journal,* vol. 46, pp. 572-590, 10, 2003.

[33] D. Elliott, W. F. Helsen and R. Chua, "A Century Later: Woodworth's (1899) Two-Component Model of Goal Directed Aiming," *Psychol. Bull.,* vol. 127, pp. 342-357, 2001.

[34] R. A. Guzzo and M. W. Dickson, "Teams in organizations: Recent research on performance and effectiveness," *Annu. Rev. Psychol.,* vol. 47, pp. 307, 1996.

[35] D. E. Meyer, D. E. Irwin, A. M. Osman and J. Kounios, "The Dynamics of Cognition and Action: Mental Processes Inferred from Time-Accuracy Decomposition," *Psychol. Rev.,* vol. 95, pp. 183-237, 1998.

[36] R. Plamondon and A. M. Alimi, "Time/Accuracy Trade-offs in Target-directed Movements," *Behav. Brain Sci.,* vol. 20, pp. 279-303, 1997.

[37] R. D. Rogers and S. Monsell, "Costs of a Predictable Switch Between Simple Cognitive Tasks," *J. Exp. Psychol.,* vol. 124, pp. 207-231, 1995.